DOCUMENT RESUME

ED 345 512 FL 020 177

AUTHOR Wall, Dianne; Alderson, J. Charles
TITLE Examining Washback: The Sri Lankan Impact Study.
PUB DATE 92
NOTE 30p.; For a related document, see FL 020 178.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; Educational Environment; *English (Second Language); Foreign Countries; *Language Tests; Secondary Education; Second Language Instruction; Second Language Learning; *Standardized Tests; Teaching Methods; *Testing
IDENTIFIERS Impact Studies; *Sri Lanka; *Teaching to the Test

ABSTRACT

A study in Sri Lanka concerning the effects of second language tests, specifically the O-Level examination in English as a Second Language, on classroom language instruction is reported. The study investigates the phenomenon of washback or backwash, the influence of testing on instruction. It is cited as the only known research investigating washback in language education through classroom observation. The study was conducted at the secondary school level, and combined classroom observation with data from interviews, questionnaire responses, and test analyses to determine whether washback exists, to what degree it operates, and whether it is a positive or negative force in this educational context. The report gives background information on the project; discusses the characteristics of positive and negative washback in terms of instructional content, instructional methods, and techniques, and assessment and presents the results of two rounds of classroom observation. It is concluded that washback occurred in both positive and negative forms, to some degree, in teaching content, but not in methodology. Evidence of washback, both positive and negative, on the way teachers and local education officers design tests was also found. An 11-item bibliography is appended. Further research is recommended. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Examining Washback: The Sri Lankan Impact Study

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Charles Alderson.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

by

Dianne Wall and J. Charles Alderson
Lancaster University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED345512

This paper is to be read in conjunction with another paper presented in this symposium: 'Does Washback Exist?' by Alderson and Wall. The main purposes of that paper were to examine the concept of testing 'washback' and to review the washback studies which have been conducted in language education. All of the studies that were mentioned (Wesdorp 1982, Hughes, 1988 and Khaniya 1990) offered suggestions about what washback might look like, and, in the cases of Hughes and Khaniya, claims that its presence had been established in given educational settings; however, the conclusions they put forward were based on questionnaire results or test scores rather than direct observation of what was happening in classrooms. Whilst questioning the usefulness of concepts like 'washback validity', the authors endorse Morrow's view (1986) that testing researchers should carry their studies into the classroom' in order to observe the effect of their tests in action'. They cite the Sri Lankan O-Level Evaluation Project as the only study that they have identified to date which attempts to investigate washback in language education by observing what actually takes place in classroom teaching.

The purpose of the present paper, then, is to describe the work being done in Sri Lanka to investigate the impact of an examination which was consciously intended to provide a 'lever for change' (Pearson, 1988). The paper will present a short background to the Project, discuss what positive and negative washback would look like, and then present the findings of two rounds of classroom observations.

The educational setting and the role of the examination

English is a second language in Sri Lanka, and one on which the country depends for various internal trade and social purposes and for conducting business with the outside world, Students study English from Year 3 to Year 11, and, as with most other subjects in the curriculum, they must sit an examination ('the O-Level') at the end of their 11th year. Their grades on the complete set of O-Levels will determine whether they will be allowed to continue on into pre-university courses or whether, if they do not continue, they

FL 020 177

BEST COPY AVAILABLE

will be eligible for desirable employment in the job market. Competition is intense for the few places available in higher education and for the limited number of jobs, so a student's O-Level grades, particularly in English, are among the most important in his/her academic career.

Unfortunately, in the decades of the 60s and 70s very few students managed to pass the O-Level English exam (the passing mark of 35/100 was usually attained by only 20% of the population). Even those who did pass were usually not adequately prepared for the situations in which they needed English. The teaching programme that they had followed was structurally based and emphasised the development of reading. Students had little opportunity to engage in everyday communication, either orally or in writing.

The 1980s brought many changes to English teaching in Sri Lanka. In response to the need for more practical English, the Ministry of Education, with the help of the British Overseas Development Administration, launched a number of textbook and teacher-training initiatives. A new textbook series was written for secondary schools, which was meant to emphasise reading and writing for a purpose and also oral skills. Pre-service and in-service training programmes were established to enable teachers to cope with the demands of the new materials. There was, however, a recognition that these innovations would not necessarily be taken seriously unless they were accompanied by a new examination which reflected the nature of the textbooks. This new examination was introduced in 1988, as the first cohort of students to go through Years 9, 10 and 11 of the textbook series were finishing their studies.

The O-Level examination and the textbook

The relationship between the O-Level exam and the textbook was quite explicit. The exam was meant to reinforce the work that the textbook writers had started. Indeed, the examination team needed to establish a syllabus for the textbook series, since none had existed 'a priori'; this involved an inspection of the textbooks and the drawing up a set of test specifications in consultation with the textbook writers. Inevitably the 'syllabus' that resulted was much greater than could be covered in a single exam. The exam team have therefore deliberately changed the exam with each new administration. This means that as the years go by all parts of the syllabus will be sampled. It a so means that teachers cannot rely on the same language structures and task types appearing each time, which, in theory, obliges them to cover the whole textbook series rather than engage in 'question spotting' and coaching.

The O-Level examination was originally meant to cover all four skill areas. Reading and writing were to be tested in a 'final exam' at the end of Year 11, and listening and speaking

were to be tested by means of continuous assessment tasks throughout Years 9, 10 and 11. However, the continuous assessment programme, which was introduced in most subjects in the curriculum, was eventually dropped due to practical and political difficulties. What remained was the final examination, which was meant to be administered for the first time in 1988.

The Sri Lankan O-Level Evaluation Project

Lancaster University was commissioned in mid-1988 to carry out an evaluation of the examination and its impact on classrooms. Our terms of reference included investigating the validity and the reliability of the final exam, and measuring the 'washback' on classroom teaching of this new means of assessment.

Our evaluation of the exam itself is reported elsewhere (Alderson and Wall, 1989), and so will not be discussed in this paper. Suffice it to say that the exam, despite a few 'teething problems', was considered to be a valid and fairly reliable means of testing what students were supposed to have learned in their last 3 years of English study. The remainder of this paper will be devoted to what came to be known as 'The Impact Study', whose purpose was to determine whether the O-Level examination was having any influence on the way that English language teaching was carried out in Sri Lankan secondary schools. The study is unique in that it attempts to combine classroom observation data with more traditional forms of data (interviews, questionnaires, and test analyses) to arrive at conclusions about whether washback exists, to what degree, and whether it is a positive or negative force in the educational context in question.

What would washback look like?

As discussed above,the new examination was meant to reinforce innovations that were introduced in the textbook series from 1986 onwards. These included the development of language skills rather than just a knowledge of grammar, and the ability to understand or produce texts which were deemed to be relevant to the majority of Sri Lankans. The reading skills that were to be developed included skimming and scanning, deducing the meaning of unknown words, picking out the main idea from supporting detail, understanding the communicative function or value of sentences, etc. (**Guidelines**:1). The writing skills included planning and organising information, giving information explicitly, transferring information from pictures to reports, and so on. The texts that were to be dealt with ranged from short messages to informative academic texts in the case of reading, and application forms to quite lengthy reports in the case of writing. Grammar was also taught but this was meant to be minimalised; while its importancc was recognised by the textbook designers and

teachers its treatment in the textbook series was neither frequent nor systematic.

The textbook writers also hoped to encourage innovations in methodology. Their main aim was to convince teachers to be less dominant, and to allow students to work individually and in pairs and small groups. They also gave suggestions about how various types of material should be dealt with in the classroom. These, and other essentials of the 'communicative methodology', were spelt out in the Teacher's Guide to Year 7, the first textbook in the series.

Given that the exam was intended to reinforce the changes suggested by the textbooks, it seemed natural to conclude that whether or not teachers were taking these innovations seriously before the exam came into being, they would strive to take them seriously after it had been in place for a while.

It should be clear by now that since the examination was supposed to be closely modeled on the textbook, one of the main problems the Impact Study would face would be disentangling the influence of the examination from that of the textbook. If teachers were using the textbook, then this would doubtless be a good thing, but it would be difficult to know whether this was because of the influence of the exam, or because this was what they were meant to be doing in the first place. We reasoned that as long as there was no conflict in the aims, activities, or criteria used to judge students' performance, the exam would have succeeded in reinforcing what the textbook meant to achieve. In an ideal world, this working together, or at least not working apart, of the exam and the textbook would have produced the following results:

1. Content of teaching

   Teachers would be 'teaching the textbook', because they would realise that any of the text types or tasks represented therein might appear on the final exam.

   They would not be giving more emphasis to any one skill, than the textbook gave it, because the weighting of the exam would reflect the weighting of the textbook.

2. Method of teaching

   Teachers would be using the general approach and the techniques suggested by the Teacher's Guides in the textbook series, as these would provide an efficient means of developing the skills that would be assessed on the exam.

3. Ways of assessing

Teachers would be writing tests that would mirror the tasks given in the textbook, because these would also be the kinds of tasks that would appear on the exam.

When marking their students' work, teachers would use the criteria laid down in the textbook, which would also be the criteria used by examiners when marking the O-Level exam.

When teachers were asked why they were teaching and marking the way they were they would reply that they were guided by both the textbook and the exam.

If this situation had existed it might have been difficult to separate 'positive washback' from the natural influence of the textbook. We believed, however, that it might not be as difficult to identify 'negative washback', where the exam, in spite of the good intentions surrounding its introduction, had a restraining or distorting influence on what was being taught and how. Once the continuous assessment programme was cancelled it seemed likely that the final exam, which tested only written skills, might encourage teachers to pay less attention to listening and speaking. In addition to this obvious 'danger' there were other ways in which the examination could work against the aims of the textbook if it did not succeed in reflecting the textbook as fully as it aimed to. The 'worst-case scenario' would look like this:

1. Content of teaching

Teachers would tend not to teach the whole textbook because they would realise that some skills, namely listening and speaking, were not assessed and that it was more useful to spend limited class time practising reading and writing.

Even when teaching reading and writing teachers might begin to neglect some kinds of text types or activities, feeling that these never appeared on the exam and were therefore not worth spending time on.

Teachers might abandon the use of the textbook altogether, and begin to use other materials which would have a more obvious relation to the exam. Amongst these materials might be teacher-designed materials, past examination papers or publications designed to help students to prepare for the exam. (The only publication of this sort that we knew of at the time was the official **Guidelines** booklet, which informed teachers of the possible content and format of the exam and gave sample test items. Valuable as this publication was, there was a possiblity that if it was used on its own it would

produced a 'narrowing of the curriculum' to match the exam. (Smith 1991)).

2. Methodology

Teachers would use whatever methodology they felt most expedient to help them to prepare their students for the exam. Some aspects of the approach encouraged by the textbook might be sacrificed if the teachers felt that these were not efficient means of preparing the students.

3. Assessing the students

Teachers would write tests which would mirror the tasks in past examination papers rather than the tasks given in the textbook. They would spend time adapting questions, or would simply lift them, either from past papers or from publications designed to prepare students for the exam.

Teachers would adopt the means of assessment used by the exam and would ignore advice in the textbook which went against this way of marking.

When teachers were asked why they were teaching and marking the way they were they would reply that they were driven by the exam rather than by the textbook.

We did not set out or expect to find either totally positive or totally negative washback. In fact, it would have been surprising to have found either, given that reality is more complicated than the best-case and the worst case scenarios suggest. What we did not know at the time was just how difficult it would be to determine whether washback had occurred at all, and to decide, if it had not appeared, whether this was because there was no such thing or because there were conditions in the educational setting that were preventing such a thing from getting through. The Washback Hypothesis, in most of the forms we knew of (Alderson and Wall 1992), suggested that a test on its own would make all the difference. If it was a 'good' test (i.e. if it reflected the aims of the syllabus, and its content and method) then it would produce positive washback; if it was a 'bad' test (if it did not) then it would produce negative washback. What was not mentioned in any of the formulations of the Washback Hypothesis were the other factors that might also contribute to what teaching will look like: Do the teachers understand the philosophy/approach of the textbook? Are they prepared to accept this? Are they able (intellectually and practically) to implement the new ideas? Are they aware of the nature of the exam? Are they willing to go along with its demands? Are they able (again, both intellectually and practically) to prepare their students for what is to come?

The aims of the Impact Study

As mentioned earlier, the washback studies by Wesdorp, Hughes and Khaniya depended on questionnaire and test results. What they did not offer was any description of what went on in classrooms before or after the introduction of the new examination, or any discussion of how what was going on in classrooms was or was not affected by the particular new examination.

We hoped in our study to learn something about the way teaching was conducted before the examination was introduced and the way it was conducted for several years afterwards. We were interested in finding out what teachers were doing and why they were doing what they were doing, as well as gathering opinions from trained observers about what was being done. We would also be analysing interviews, questionnaires, classroom materials and test results, but we hoped that by observing classrooms as well we could add an extra dimension to what was already known about washback.

Method

Baseline Studies

Before anyone could determine whether the new O-Level examination was having an effect on teaching, it was necessary to find out what teaching was like before the introduction of the exam. A series of baseline studies was carried out in 1988, about 6 months before the exam was meant to be held for the first time. In one of these studies members of the examination team carried out observations in a small but representative sample of schools. The observers were interested in seeing how teachers handled the textbooks, which by that time had been in use for several years, and in finding out how teachers viewed their own teaching and the influences upon it. The teachers tended to claim that they had begun using a 'communicative methodology' once they had received the new textbooks, but the observations indicated that this was not the case. Most classes were very formal, students spent much time listening to the teacher or practising language form rather than using the language, and many teachers did not seem to have clear objectives for their teaching. Teachers had little understanding of what the new exam would be like, which was natural since the official exam support material was only just beginning to arrive in their schools. The observers' reports indicated that teaching could not yet have been influenced by the exam,

Observation programme: 1990-1991

At the core of this programme were 7 teachers, based in 5 different parts of Sri Lanka, who had agreed to act as observers for the Impact Study. These teachers underwent a training programme which included discussions about the teaching materials and the philosophy behind them, discussions and analysis of the new examination, discussions about examination washback and how they would recognise this in classrooms, and practise in doing classroom observations. Each of the observers agreed to visit 7 schools in their own areas over the course of two years. The observers would fill in observation schedules for each classroom they visited, conduct interviews with the teachers they had observed, and record their own opinions about what they had seen and heard.

They would then send completed schedules to Lancaster. These would be analysed and observers would receive feedback and instructions for the next round of observations. It was planned that there would be 6 rounds of observations in all, one a term for each of 2 academic years.

Round 1, which took place in the first term of the 1990 academic year (the 'year' had started 4 months late as a result of political difficulties throughout the country) served as a pilot round, which helped us to see the problems in the instruments and procedures that had been created for the study. Round 2, held four months later and about three months before the 1990 exam was to be given, yielded useful information and gave us further insights into what we should be looking for. Round 3, which took place a month before the exam, produced data for only 14 schools. Many of the schools in Sri Lanka stop giving classes before the O-Level exams so that students can do independent study. It had not been realised before Round 3 just how early many schools ceased teaching, and when observers appeared at the schools to visit classrooms they found that there were no classes to observe. The fact that two-thirds of the English language classrooms had 'dissolved' was our first real indication of how much influence any examination was bound to have in this particular setting. A review of the English language classrooms which were in session showed that 10 of them were doing some sort of 'examination practice'. This gave us a hint of the influence of this particular exam, at least at the end of the academic year, but there were not enough schools represented in this round to allow us to make generalisations about classrooms all over the country.

Rounds 4 through 6 took place in 1991, the first 'normal' academic year that much of Sri Lanka had experienced since before the introduction of the exam. Intense fighting in some parts of the country means that there are school closures and other disruptions even now, but in the areas covered by our study schools were in session for a full three terms before the exam was given. The sample size fluctuated from round to round, due to the departure of several of the original

observers, the arrival of four new team members, and the difficulties that all observers had in being released from their regular teaching duties to carry out the research. At its largest the Impact Study encompassed 64 schools; at its smallest (excluding Round 3) it encompassed 36. There were 18 schools that were observed at every stage in the programme.

It is not possible in this short paper to discuss the findings of the earlier rounds, except to say that they matched one another and gave us a good indication of what was to appear later in the study. The rest of this paper will be based on the most recent observations: Round 5, the largest round and, we believe, a 'typical' one, and Round 6, when much formal 'exam preparation' was taking place. There will also be references to interviews with many of the teachers who were observed during the two years, to see if they could help us to understand better what they had been doing and why.

What did classes look like?

Round 5: a 'typical' round

In Round 5 (June/July 1991, 5 months before the exam) observers visited 64 Year 11 classrooms. Of these approximately three-quarters were using the textbook and approxmately a quarter were using other sorts of materials in class. The observers wrote detailed descriptions of what happened in each classroom and filled in checklists recording the types of texts that were used and the types of activities that took place. They also recorded their views of the effectiveness of the particular lesson, and whether the teaching in the lesson might have been influenced by the examination. An analyst in Lancaster also made judgements about whether the teaching might be influenced by the exam, and if so, how. It is important to note that the observers and the analyst were not asked to say whether exam impact definitely existed: they could not have known this without analysing data from other sources, particularly interviews with teachers. It is because of this uncertainty that phrases such as 'There might be impact' and 'There was a possibility of impact' occur so frequently in the rest of this paper.

Table 1 gives the analyst's initial view about possible exam impact on these classes:

Table 1: Analyst's initial view of impact, Round 5'

	There might be exam impact	There might not be exam impact
Teachers using the textbook	31 teachers (70%)	13 teachers (30%)
Teachers not using textbook	17 teachers (100%)	-

' Figures available for 61 classes at time of writing.

Classes where the textbook was being used (n=44)

There was a possibility of exam influence in about 70% of the classes where the textbook was being used, and little evidence of impact in the other 30%. In all cases the resemblance resided in the content of the classes: in the reading lessons this meant the types of passages that were being studied and the types of questions that were being asked, and in the writing lessons, the type of text that students were asked to produce. If the lessons emphasised listening or speaking skills then they were not thought to be influenced by the exam. It was possible, however, for teachers to turn what were designed to be oral lessons into lessons where written skills were emphasised instead: this happened in the case of two role plays where the teachers asked the students to answer comprehension questions, and another role play where the teacher turned the text into a C-test passage.

Of course, the fact that the content of lessons resembled the exam did not mean that the teachers were designing lessons to match the exam. In almost half the classes the teachers were taking the content (text and question types) straight from the textbook. They had added nothing of their own to make the content more like the exam. When these teachers were asked why they had chosen their content for the day, two-thirds replied that they were doing what came next in line in the textbook. Only 2 replied that they were intending to prepare their students for the exam.

In the other half of the classes where the content resembled the exam, the teachers were using the texts that appeared in the textbook but adding questions or other tasks. These questions/tasks were of the sort that might appear on the exam. The additions mainly took the form of short-answer questions written on the blackboard (12 cases out of 17), true/false question, a C-test passage and a writing task.

It is natural for teachers to wish to check the comprehension of their students by asking them questions, especially if the textbooks that they are using and/or their training have not provided them with other means of making sure that their students understand. It is therefore not necessarily the case that the addition of comprehension questions indicates the influence of the exam. Indeed, only a quarter of the teachers who added questions reported that they were attempting to prepare their students for the exam on the day of the observation; the others, with 1 or 2 exceptions, reported that they were teaching the lesson that was next in line. The fact that so many teachers claimed to be teaching what came next in the textbook suggests that there was no exam impact in these cases. While it is always possible that the teachers would not have been teaching the next lesson if they had not believed it relevant, the observations offer no evidence to prove this.

There were two other factors which might have indicated the presence of exam impact on the content: the amount of attention paid to each skill area and the references that

teachers made to the examination in their lessons. The exercises that the teachers had chosen for the day were meant to be focusing on the following areas:

Reading	52% of the exercises
Writing	17%
Listening	5%
Speaking	17%
Language form	10%

The number of exercises which were meant to be developing reading skills is far greater that the number for other skills, especially listening. This might have supported the idea of examination impact since roughly half the examination is devoted to reading; however, it does not explain why writing, which also accounts for about one half of the examination, was the focus of so few exercises. A possible answer lies in an inspection of the Year 11 textbook, which was being used by almost all of the teachers on the day of the observations. A tally was made of the exercises devoted to each skill from Lesson 4 onwards:

Reading	40%
Writing	20%
Listening	10%
Speaking	25%
Language form	5%

By comparing this tally with the previous one it can be see that there was some difference in the amount of attention that the teachers were meant to be paying to each skill and the amount of coverage given to the skills in the textbook. (N.B. The phrase 'meant to be' must be used here because teachers frequently focus on language form during reading lessons, have students read aloud during reading and writing lessons, correct spelling during listening lessons etc.) Reading seems to be claiming more attention than it should, perhaps at the expense of listening and speaking. Although most of the teachers reported that they were teaching the content that was next in line, several who were teaching reading reported that they were preparing their students for the exam. This may explain some of the discrepancy between the two sets of figures ; however, the general impression that teachers were devoting more time to reading than to the oral skills does not seem so surprising given the minor attention that the textbook pays to oral skills (especially listening) in the first place.

The other aspect of content which was analysed was the number and type of reference made to the exam during the lesson. Only 6% of all the teachers using the textbook on the day of the observation made reference to the exam: when they did it

was only to point out that a certain text type or type of question had appeared on the exam before or might appear in future.

There is little evidence, then, that the examination was causing teachers to modify the content of lessons if they were using the textbook. There may have a small amount of negative washback in that some teachers chose to handle reading lessons rather than the lessons that were next in line; however, this did not involve more than 2 or 3 teachers out of 44 and should not be considered significant.

There is no evidence that the examination was affecting the teachers' methodology, either positively or negatively. The official exam support documents list 'Skimming to obtain the gist' and 'Finding specific information' as two of the skills which are likely to be tested, and the fact that the students have to answer questions on many reading passages, some of them fairly long, would suggest that reading quickly and disregarding irrelevant detail would be useful skills to develop in the classroom. The observers' reports indicate that these sorts of activities rarely take place. Teachers do not train their students to read selectively; in fact, many teachers seem to believe that students need to understand all the words and the grammar of every passage, regardless of the nature of the questions given to assess their comprehension.

The Teacher's Guide to Years 10/11 does not give any advice on developing skimming and scanning abilities. It often provides 'pre-reading' and 'scanning' questions, but it assumes that teachers will know how to use these questions to their best advantage (timing students, giving them hints as to how to find key information, etc.). The 'Finding Out' procedure is alluded to frequently, but this procedure, which is explained in the Teacher's Guide to Year 7, recommends that teachers should teach reading by giving background to the topic, clarifying difficult structures or vocabulary, reading the passage aloud, having the students read it silently for 5 minutes, reading it aloud again, and getting the students to read it aloud - all before asking the students to answer any content questions! Although this method might be suitable for students at the start of their secondary school studies, it seems less appropriate for O-Level students, who must read longer and more complex passages more efficiently.

What is ironic, however, is that many teachers provide even more 'support' than is recommended by the earlier Teacher's Guide, to the point of 'spoon-feeding'. Teachers in all the rounds have been observed explaining all the difficult words of the passage (cf Teacher's Guide: 'The pupil's own skill of guessing intelligently from context and relating mutually explanatory parts of the passage MUST be given scope'), dissecting passages sentence by sentence (cf 'They must learn to focus their attention on the GENERAL message in the first instance and not on minor details') and doing considerable amounts of explaining in the first language (cf 'Don't kill

their interest by giving them everything "on a plate" in advance'). Students are often required to read at the pace of the group rather than at their own pace. It is worth emphasising that such teachers are not only not being influenced by the exam, but they are also not being influenced by the 'philosophy' of the textbook series. There are many factors which could explain this: inadequate or outdated teacher training, inadequate distribution of Teacher's Guides, poor preparation of students in earlier years of the curriculum, over-ambitiousness of the syllabus etc. These and other possible explanations will eventually be explored further in our Final Report, by references to individual and group interviews of teachers, and questionnaire surveys of teachers and teacher advisers.

There is no evidence that the methodology used in the teaching of writing has been influenced by the exam. An inspection of the tasks given in the exam and the criteria for judging students' performance indicate that relevance of ideas and a certain sensitivity to audience are important factors in writing, but these are matters of content and do not suggest or require that any particular methodology. What the observers' reports show is that the teachers of writing, as was the case with the teachers of reading, tended to disregard the advice given in the Teacher's Guides. The most obvious difference between the way they taught and the way they were advised to teach was that they tended to prepare students for writing by doing all the support work themselves rather than allowing the students to work together and help each other. However, this altering of the recommended methodology could be the result of many factors (ignorance of suitable techniques and fear of noise, to name but two), and cannot be traced to the influence of the exam.

To summarise, then, there is no evidence that the exam is having any impact on methodology. In fact, the observations suggest that the way that teachers present their content and get students to practise the language runs contrary not only to the aims of the exam but also to some of the basic principles of the textbook series. There are many reasons why this may be happening, but the most obvious one emerging from interviews is that many teachers do not understand what the exam is really testing or what the textbook is teaching. (See the section on 'The need for complementary data' below for further discussion of this point.)

Lessons where the textbook was not used (n=17)

As might have been expected it was easier to find exam impact in lessons where the textbook was not being used. All of the teachers except one said that they intended to use their lessons for 'examination preparation': two-thirds of these referred specifically to the O-Level examination, and one-third referred to teacher-made tests, which we know from other studies (Wall 1991) often resemble the O-Level examination.

There were 17 classes where the textbook was not being used. Of these 7 focused on reading, 7 on writing, and 3 on language knowledge. None of the classes focused on listening or speaking. In fact, there was almost no oral work to be seen, apart from the students listening to the teacher and answering the teacher's questions.

The resemblance that was seen between the reading classes and the exam lay in the types of passages being used and the types of tasks that students had to complete. The passages and tasks were taken straight from past examination papers or from commercial exam preparation books in all cases but one. The passages included newspaper accounts, advertisements, forms, letters and descriptions, and the activity types involved jumbled sentences, and multiple choice, true/false and short answer questions based on information that was directly stated in the passage.

The resemblance between the writing classes and the exam resided in the types of text the students had to produce. These were taken straight from past papers, official exam support material, or commercial examination preparation books in all cases but two. The tasks included filling in forms, writing diary entries, argumentative writing, descriptive writing, invitations and poster captions.

The resemblance between the language knowledge lessons and the exam also resided in the type of task being done. The tasks were designed by teachers but the task types had all appeared on past exams: gap-filling and grammatical transformation.

The methodology of these lessons, however, was not so obviously tied to the examination. This was especially true in the case of reading lessons. Roughly the same pattern was followed in all of the classes:

Teacher writes passage on board
Students copy passage into copybooks
Teacher and/or student reads passage aloud.

Teacher writes (occasionally dictates) questions
Students copy questions into copybooks.

Teacher dwells on instructions, often using L1.

Students take much time to answer questions on own (although occasionally they work in lockstep, question by question).

Teacher asks for answers and students give them.
Teacher and students discuss incorrect answers (sometimes).

Teacher asks how many students have got all the questions right, all but one, all but two etc.

Though it was not as common as in teaching using the textbook, teachers occasionally dissected texts before they asked students to answer the questions - explaining or translating the difficult words, paraphrasing or translating difficult sentences.

Interestingly, though, this explaining or translating process was often lacking when it was most needed, when students had given incorrect answers and needed help to find out how they had gone wrong. The students' answers were discussed in about 60% of the lessons but in the remainder they were merely accepted by the teacher or rejected.

This methodology obviously eliminates any possibility of skimming or scanning. Students usually read through the passage several times (as the teacher was writing it on the board, as they were copying it into their copybooks, and as they heard it read aloud or read it aloud themselves) before they read and copy the questions. There was therefore no opportunity for them to practise rapid selective reading. When the students were allowed to work on their own they often worked with no strict time limit, so they did not get used to the idea of having to read quickly. When the students worked in lockstep, some did not have a chance to practise reading at all as the students who worked more quickly were often requested by the teacher to supply answers to the whole group. The cases where the teacher explained everything before asking the students to answer the questions were worrisome, but equally worrisome were the cases where the teacher either accepted or rejected the students' answers without discussion. A student could learn that he/she had responded incorrectly, but not be able to find out why.

The pattern was less rigid in writing lessons but nonetheless visible:

Teacher writes rubric on blackboard.
Students copy rubric into copybooks.
Teachers spends some time explaining instructions, often using L1.

Teacher asks students questions about the task.
Teacher either puts key words or full sentences on blackboard. (On one occasion teacher dictated full sentences.)

Students write individually.

Correction:

Students read what they have written to rest of the group. Teacher corrects the student who is reading and the other students try to correct themselves, or

Teacher walks around correcting as he/she goes, or

Students take copybooks to teacher for correction.

In one case students were allowed to brainstorm within a small group. The group then stayed together to do the writing themselves.

In several cases no correction was done because the writing itself took up the whole class session.

In no case were the criteria that were being used to correct written work made explicit to the students.

It can therefore be seen that although the content of the writing lessons matched the content of past examination papers, the methodology bore little resemblance to what students would have to do when sitting the exam. The biggest difference seems to be in the amount of support the teachers gave the students, including writing out models for students to copy. Although it is the teachers' responsibility to help the student (when teaching, as opposed to when testing), excessive help can sometimes kill the students' initiative or even cause frustration. It might be more effective to allow the students to come up with their own ideas rather than supplying them. It would also be useful if the students were made aware of the criteria that will be used to judge them on the O-Level examination. Some sort of group effort at correction might also be worthwhile. Obviously teachers do not have enough time to correct every student's work; however, the observations showed that students who tried to correct their own work as another student's work was being corrected orally sometime did not hear the corrections and often did not heed them.

In summary, the Round 5 observations suggested the following:

1. If teachers were using the textbooks in their lessons, they either took the content (texts and practice activities) straight from the book or modified it by adding questions or a task. There was no evidence, however, that they made these modifications because of the exam, although they were all in the direction of the exam.

2. If the teachers were not using the textbook the content of their lessons was very much influenced by the exam.

3. The methodology that the teachers were using, whether they were teaching from the textbook or not, showed no impact from the exam. Indeed, if often showed little impact even from the Teacher's Guides to the textbooks.

Round 6: 'Exam preparation'

In Round 6 (October/November 1991, approximately one month before the examination), observers visited 41 Year 11 classrooms. Of these, 29% were using the textbook and 71% were using other sorts of materials in classes. Table 2 gives the analyst's initial view of the possibility of exam impact on these classes:

Table 2: Analyst's initial view of impact, Round 6

	There might be exam impact	There might not be exam impact
Teachers using the textbook	11 teachers (92%)	1 teacher (8%)
Teachers not using the textbook	29 teachers (100%)	-

As in Round 5, it seems clear that teachers who bring in supplementary materials tend to give lessons which are very similar to the examination.

Classes where the textbook was being used (n=12)

There was a resemblance between the content of the classes and the content of the examination in 11 out of 12 classes. In only 2 cases, however, did teachers change the content of the textbook lesson they were working from, and, as in Round 5, this meant adding exercises (a set of short-answer questions in one lesson, and the gapping of a text in another). The fact that teachers added exercises to a lesson does not necessarily indicate that they are influenced by the exam.

Most of the teachers who were teaching from the textbook were working on one of the last three units of Book 11. A tally

showing the relative amount of attention paid to each skill in these three chapters and the proportion of the teachers who were working on exercises which were supposed to develop these skills shows the following:

	Percentage of exercises in the textbook devoted to skill	Percentage of teachers who had chosen exercises to develop skill
Reading	32%	58%
Writing	27%	33%
Listening	7%	0%
Speaking	27%	8%

These figures show that teachers were paying less attention to the oral skills than the textbook itself would have had them do, and more attention to reading. This emphasis on reading may indicate examination impact; however, this is difficult to prove given that all but 2 of the teachers reported that they were teaching the lesson which came next in line in their textbooks.

There were very few references to the exam in the lessons. There were only 2 or 3 references each to text types and activity types which might appear on the exam.

There seems to be little evidence that the exam is having any independent impact on the content of the lessons. The methodology was very much the same as was found in Round 5, and the comments made about the apparent inappropriateness of the methodology when compared to the goals of the exam and the textbook also apply here.

Lessons where the textbook was not used (n=29)

As in Round 5 it was easier to find exam impact when teachers were not using the official textbook. All of the teachers who responded to a question concerning the purpose of the day's lesson reported that they were doing 'exam preparation', and all but one of these referred specifically to the O-Level exam rather than to internal year-end tests.

The teachers either used teacher-designed materials (about a quarter of the sample) or commercial publications which are designed to help teachers and students to prepare for the exam (about half the sample). Only 13% were using past examination

papers, and only 1 teacher was using official exam support material. It is interesting to note the large difference between the number of teachers using commercially-produced materials and the number using official exam support materials. Although the latter are well written and very informative, they do not seem to be used very often in classrooms. This may be a problem of distribution (only half the teachers either owned or had access to either of the two support booklets) or it may be that the commercially-produced materials (which are available in the urban areas, for a price which is equivalent to half the daily earnings of the typical secondary school teacher) hold some other attraction which we have not yet uncovered.

The skills that the teachers were concentrating on during Round 6 were:

Reading	15 teachers	52% of valid sample
Writing	9 teachers	38%
Language form	5 teachers	17%

There was no attention paid to either listening or speaking. The students listened only to their teacher's reading aloud or explanations of lessons. The most common form of 'speaking' was answering questions asked by the teacher, although occasionally some students were asked to read aloud.

In the reading classes the teachers presented a variety of text types, the most common being informative academic or semi-academic texts (e.g. excerpts from encyclopedias or textbooks) and descriptions of persons, places or things. All of the text types appear on the exam, and the two most common ones appear frequently. Students were asked to answer several different types of questions, all of which had appeared on past papers, but no one sort predominated.

In the writing classes students were practising the filling in of application forms in all the classes but 2. This is a type of writing that has appeared on every version of the new exam.

In the language form classes students were practising grammar transformation exercises in all classes but 1.

The examination was obviously having an effect on the content of all these classes, with the most obvious effect being the virtual disappearance of listening and speaking, and the attention being given to certain types of reading passages and to form-filling and transformation exercises.

The Round 6 exam preparation classes followed the same general patterns for methodology as in Round 5. In approximately half the reading classes the students did not have copies of the passages they were supposed to study. This meant that large amounts of class time were spent on writing: the teacher transferring a text and questions from a past paper or a commercial publication onto the blackboard, and the students

copying from the board into their exercise books. Sometimes the students spent so much time copying that there was little time left for answering the questions or for checking whether the answers were correct. An observer wrote the following description of one such lesson:

8.35 Teacher started writing the text on the blackboard: an advertisement calling for applications for trainee accounts clerks.

8.45 Teacher and students still writing or copying the text.

8.50 Teacher wrote comprehension questions on the blackboard. 6 questions, e.g.

1. When is the closing date?
2. Can you forward typewritten applications?

8.55 Teacher asked students to answer the questions in their exercise books.

9.00 Students still writing.

9.10 The period over.

In the other half of the reading classes the teachers either borrowed class sets of books which came earlier in the textbook series (students have to return their books to the school at the end of each year), or collected money from the students to pay for photocopies of past papers, or asked the students to buy copies of the commercial publications. Photocopying is much less expensive than getting students to buy books; however, both options are beyond the means of most families. The observers reported only one teacher who had found a way around the problem of providing supplementary texts in poor areas: she brought in authentic texts from newspapers and distributed them amongst the students, allowing them to read and answer questions at their own pace and then providing answers for each student. This is the kind of activity that Teacher's Guides to the textbook and the examination support materials should be providing, but unfortunately still are not.

In classes where the students were concentrating on shorter texts or where they had their own copies of texts, the teachers might set three or four 'model questions' in a single period. Here the pattern was very tedious: students copied the text, answered the questions, gave their answers to the teacher, found out whether they are wright or wrong, and then passed onto the next exercise. There seemed to be less

attention paid to clarifying instructions and to discussing the answers in Round 6 than in Round 5.

The methodology for writing classes and for language form classes was much the same as for Round 5. Again, if texts and questions were short some groups could manage to get through more than one exercise in a period. However, it would be difficult to try to deal with any of the longer exercises that are found toward the middle and the end of the O-Level exam because there would be too much to copy. This meant that students from poorer families and in schools with fewer resources were not always able to engage in certain types of exam practice because it took too much time to copy texts from one place to another.

What are the observations showing us?

These two rounds of observations, in combination with previous ones, seem to suggest that

Most teachers follow the textbook during normal teaching terms. They work their way through the materials, unit by unit, exercise by exercise. This may be because they believe they have to 'cover the book' so that their students will do well on the exam, or it may be that they do not know any better.

Less attention is paid in Year 11 to the development of listening and speaking skills than to written skills, even in normal teaching terms. This may be the effect of the exam; however, it may also be due to the fact that the textbook pays less attention to these skills, or that teachers do not know how to teach listening and speaking.

There is little visible exam impact on the content of reading, writing and grammar lessons if teachers are using the textbook. Teachers occasionally add questions or tasks to the day's lesson, but this may be to compensate for a lack of suitable exercises in the textbooks and not because of the exam. It is important to note, however, that the changes they do make are always changes in the direction of the exam.

The third term of the academic year is very different from the normal teaching terms. It is clear that there is a 'narrowing of the curriculum' as teachers finish or abandon their textbooks and begin intensive work with past papers and commercial publications to prepare their students for the exam. At this point there is obvious exam impact on the content of the teaching.

There is no relationship between the methodology that teachers use, whatever the time of year, and the methodology that might be most suitable for students to use when sitting the exam. (It is worth noting here that

when the observers were asked to judge the effectiveness of the classes they had visited they judged them 'effective' in fewer than half the cases (40% in Round 5, and 45% in Round 6. They were not convinced that many teachers understood the basic principles of the textbook they were using or that they were in command of communicative teaching techniques.)

These findings are interesting, but it is clear that observations on their own cannot give the full picture of what is happening in classrooms, and that it is necessary to take advantage of other means of data collections. The next and final section of this paper will discuss the importance of collecting data from a variety of sources.

The need for complementary data

Although the observations provided us with many insights into the relationship between teaching and the examination, they also left us with questions which no amount of observation could answer. These included queries like the following:

Many teachers reported that they were teaching the lesson that was next in line, even in Round 6. Why were they doing this?

Some teachers reported that they had selected certain material in the textbook in order to prepare their students for the exam. What kinds of material had they skipped over, and why?

Were there any kind of material that teachers consistently missed out?

If teachers brought in supplementary material, what skills did they hope to develop?

What was the attraction of the commercial publications that so many teachers were using in Rounds 5 and 6?

How much did the teachers really understand of the aims of the textbook series?

How much did they know about the exam?

How much influence did they feel the exam had had on the way they chose their content and methodology, and the way they designed and marked their tests?

Did they believe the exam influenced their teaching in Years 9 and 10, when the exam was still a long way off?

It was therefore useful for us to complement the classroom observations with teacher interviews, questionnaires to teachers and teacher advisors, and analyses of materials (especially tests) that teachers had prepared for classes. It is not the purpose of this paper to report the findings of all these studies (see Wall 1991, and Alderson and Wall (forthcoming) for more details); however, it is important to record that if we had not sought to 'illuminate' the findings of the observations with other sorts of data we would have missed insights like the following:

1. Many teachers believe they have to follow the textbook faithfully because the exam may test any of the content therein. Many teachers give extra classes to their students after normal school hours, on weekends and during holidays - not to work on special exam preparation material but simply to cover as many units as possible in their textbook. This implies exam impact on how teachers choose their content. (Data gathered from group interviews.)

2. Many teachers believe that they have to cover the earlier textbooks in the series equally thoroughly, especially the Year 10 book. An important factor contributing to this belief is that past exams have taken some passages straight from earlier books in the textbook series. The teachers call these 'seen passages' and believe that if their students have studied every passage in every book they have a better chance of recognising any passage which appears on the exam - even though the questions/tasks accompanying the passages may be different. This implies exam impact on how teachers choose their content. (Group interviews.)

3. A number of teachers, however, consistently skip over the listening lessons in their textbooks, because they know that listening will not be tested in the exam. Other teachers may 'do listening', but in a way that does not resemble the textbook designers' intentions. One teacher, for example, admitted that he only covers the listening lessons if the type of question that students have to answer resembles an item type that might appear in the examination for reading. Other teachers use the listening comprehension texts, which are printed in the back of the student book, as passages for reading practice. This implies exam impact on lesson content. (Group interviews)

4. Many teachers report that they continue to teach listening and speaking; however, they admit that they do not test oral skills in their classrooms. The testing of listening and speaking seems to have died out with the demise of continuous assessment. This implies examination impact on the way teachers assess their students. (Group interviews, Analysis of tests)

5. Examination preparation takes up a large amount of the teaching year. Some teachers begin giving exam preparation classes in Term 2 (note that approximately one quarter of the teachers were doing exam preparation in Round 2); most teachers report that by late October or early November they are spending much of their class time and often extra time (again, after school, weekends, public holidays) helping their students to prepare for the exam. This implies exam impact on content. (Group interviews.)

6. Many teachers know less about the exam than they realise. Most teachers can list the types of passages that might appear or the types of writing tasks, but they may not understand what is really being tested. Some teachers report that they teach the 'content' of the reading passages in their students' textbooks or in exam preparation books, because they believe their students will need to know facts such as the names of parts of a computer, the characteristics of certain animals, the number of radio transmitters in the United States, and so on. If their students learned these facts then they would more easily be able to understand 'seen passages', or would be able to use this information in the writing subtests. The notion of reading in order to get new information rather than confirm old information is not universal. Many teachers are also unfamiliar with the criteria that will be used to mark student writing. This lack of understanding might lead to negative washback, although the exam paper itself is not responsible. (Group interviews.)

7. A quarter of the teachers have not received any training on how to use the textbooks, and as many as 40% do not have access to some of the Teacher's Guides. This ignorance could prevent examination impact from coming through. (Individual interviews.)

8. Only one-third have received any training on how to prepare students for the exam and only half have access to copies of official examination support materials. Again, this ignorance could prevent exam impact from coming through. (Individual interviews.)

The sceptic might ask why, if we were able to find out so much from individual and group interviews and from other forms of data collection, it was necessary to go to all the expense and trouble of observing classrooms. There are three answers to this question:

1. If there had been no classroom observations, we would not have known that certain questions needed to be asked. For example, it might not have occurred to us to ask about the teachers' training and access to Teacher's

Guides if we had not observed that many of them did not seem to understand what it was they were meant to be teaching, We might not have asked about commercial publications had we not seen how often teachers use them rather than official exam support material. The list of such questions is long.

2. If there had been no observations, we might not have been able to understand some of the answers that teachers gave us, especially when they referred to other factors that helped to determine what they taught and how.

3. If we had not analysed what goes on in classrooms we would have had no choice but to believe what the teachers told us. We might not have so readily doubted the claim that 85% of the teachers made in Round 5 and 90% of the teachers in Round 6: that the examination has influenced how they chose their methodology. Classroom observations reveal that this is definitely not the case, and they have helped us to appreciate that the exam can have no impact on methodology unless the teachers understand correctly what it is the exam is testing. Interviews or questionnaires on their own would have painted a more positive picture of washback than observations will allow us to accept.

It is clear, then, that observations on their own can only reveal part of what is happening within any educational setting: the observers can see what is going on, but they may not understand all they see. The other forms of data-gathering, though, will be equally uninformative if not accompanied by an analysis of teaching. Without observations the researchers may not know all the questions they should be asking and might not understand (or doubt enough) the answers they will be given.

Earlier in this paper (page 6) we questioned whether an exam on its own could make a difference in teaching. We referred to other factors that might contribute to what teaching looks like, and stated that although they seem obvious to us now they were not so apparent when our research began. It has only been by combining observations with other types of studies that we have been able to see the following:

A considerable number of teachers do not understand the philosophy/approach of the textbook.

Some teachers are not prepared to accept all aspects of this approach.

Many are unable, or feel unable, to implement the recommended methodology.

Many are not aware of the nature of the exam - what is really being tested.

All seem willing to go along with the demands (if only they knew what they were).

Many are unable, or feel unable, to prepare their students for everything that might come.

Conclusion

We conclude by returning to our hypotheses about positive and negative washback (pages 4 and 5) and listing our findings to date, from observations and other forms of data collection.

1. Content of teaching

There is evidence of washback on the content of teaching. some of this is positive and some negative.

Evidence: Teachers' determination to cover the textbooks (mostly positive, but negative if the motivation is to memorise texts which may become 'seen passages'), more attention being paid to reading than textbook provides for (negative), less attention paid to oral skills than textbook provides for (negative), claims from teachers that they skip listening lessons because the exam does not test this skill (negative), long examination preparation period with materials reflecting content of exam and abandonment of listening and speaking (negative).

2. Methodology

There is no evidence of washback on methodology. Teachers cannot tell by looking at the exam how they should teach reading or writing, and the official exam support materials do not help on this matter. However, Teacher's Guides to the textbooks offer little advice on methodology in Year 11 and advice in earlier years seems to contradict the type of methodology that would be suitable for the aims of the syllabus agreed between textbook writers and exam design team.

3. Ways of assessing

There is evidence of washback on the way teachers and local education offices design tests. Some of this is positive and some negative.

Evidence: More attention to reading and writing than to grammar (positive), much use of item types which have appeared on the exam (positive when these have also appeared in the textbook, but negative when they have not and when certain types are over-used), much copying of passages and questions straight from past papers (negative).

There is no evidence of washback on the way that teachers mark their students' class tests and assignments. Few teachers have served as markers or received any training concerning the exam. Many have not received official exam support materials which explain the marking system However, Teacher's Guides to the textbook also lack advice in this area.

Finally, we must reiterate our belief that the Washback Hypothesis needs considerable investigation and clarification before the existence and nature of washback can be accepted. The issue is not just 'Is there washback?', but also 'What is it?'. We hope that the Sri Lanka Impact Study will contribute to the general understanding of these questions.

BIBLIOGRAPHY

Alderson, J.C. and D. Wall (1989) **The Sri Lankan O-Level Evaluation Project: First Interim Report**. Lancaster University.

Alderson, J.C. and D. Wall (1992) 'Does Washback Exist?', paper presented at the Language Testing Research Colloquium, Vancouver.

Alderson, J.C. and D. Wall (forthcoming) **The Sri Lankan O-Level Evaluation Project: Final Report.** Lancaster University.

Hughes, A. (1988) 'Introducing a needs-based test of English Language Proficiency into an English-medium university in Turkey'; in Hughes, A. (ed.) **Testing English for University Study**, ELT Documents 127, Modern English Publications, London.

Khaniya, T.R. (1990b) 'Examinations as Instruments for Educational Change: Investigating the Washback Effect of the Nepalese English Exams'. Unpublished PhD thesis, University of Edinburgh.

National Institute of Education, Ministry of Education, Sri Lanka. (19xx) **Teacher's Guides to Years 7 and 10/11**, English Every Day.

National Institute of Education, Ministry of Education, Sri Lanka. (1988) **Guidelines to the O-Level Examination for 1988 and Beyond.**

Pearson, I. (1988) 'Tests as Levers for Change' in Chamberlain, D. and Baumgardner, R. (eds.) **ESP in the Classroom: Practice and Evaluation**. ELT Documents 128. Modern English Publications

Smith, M. L. (1991) 'Put to the Test: The Effects of External Testing on Teachers', in **Educational Researcher**, Volume 20, Number 5, June-July 1991.

Wall, D. (1991) 'Measuring Washback: The Sri Lankan Evaluation Project', paper delivered at the International Association of Educational Assessment conference, Nairobi

Wesdorp, H. (1982) 'Backwash effects of language testing in primary and secondary education'. Stichting Centrum voor onderwijsonderzoek van de Universiteit van Amsterdam, Amsterdam